Comment on nhess-2022-125

Rui Jesus (Referee)

The paper describes a framework for building damage classifications after an earthquake that combines InSAR data, high-resolution building inventory data and earthquake ground shaking intensity maps. The classification is performed using two different strategies (multi-class and binary classification) and two different machine learning algorithms.

I am a researcher in the field of computer science, particularly in machine learning. Therefore, this review will focus on issues related to machine learning techniques used to solve a problem in the area of natural disasters.

The article is well structured and organised. It is not difficult to understand the main ideas of the work done and it is well written. The work described represents a very valid proposal but the original contributions are minor. Anyway, if the proposed framework proposed is well evaluated and validated, for me, it can be accepted for publication. However, as the paper is, in terms of validation, does not follow the standards and requirements of this journal.

The authors claim innovation in the use of the Multi-class damage grade classification using InSAR data. I have nothing against but the results present are weak to validate the method. I think authors should present results that clearly show the benefits of the strategy used when compared to other techniques proposed. The multi-class strategy seems to me very important but the results are weak and are not compared. Also, the selection building damage categories is a critical point, regarding the urgency of action. This should be more discussed and explained in the paper.

Deep learning techniques are been used to solve many problems in many fields and presenting good results or better than the previous methods. I found it very strange that the authors would talk about recent advances in machine learning and then not use deep learning techniques for classification.
I understand that in some situations it can be difficult to use deep learning techniques, particularly when there is not much data available for training. If that's the case, I think the authors should present results demonstrating this. Also, the data augmentation methods should be considered.